



# Critical Study: Cassam on Self-Knowledge for Humans

## Citation

Boyle, Matthew. 2015. "Critical Study: Cassam on Self-Knowledge for Humans." *European Journal of Philosophy* 23, no. 2: 337-348.

## Published Version

doi:10.1111/ejop.12117

## Permanent link

<http://nrs.harvard.edu/urn-3:HUL.InstRepos:22557392>

## Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Open Access Policy Articles, as set forth at <http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#OAP>

## Share Your Story

The Harvard community has made this article openly available.  
Please share how this access benefits you. [Submit a story](#).

[Accessibility](#)

## Critical Study: Cassam on Self-Knowledge for Humans<sup>1</sup>

Matthew Boyle, Harvard University

Not so long ago, in the late 20<sup>th</sup> century, philosophers of mind generally assumed that we have some sort of privileged access to, or at any rate some special authority in speaking about, our own mental states. Most did not assume that this privilege extends to every aspect of our mental lives, but most did assume that we are at least in a specially favorable position to know our own thoughts and feelings, on the one hand, and our own standing attitudes such as belief, desire, and intention, on the other. Moreover, late twentieth century philosophers typically assumed that this privilege reflects some difference of principle in the way we can gain knowledge of these topics. The problem was to explain this difference of principle.<sup>2</sup>

Times change, however, and there are now a growing number of philosophers who question the datum this tradition set out to explain. Quassim Cassam's *Self-Knowledge for Humans* (2014) is a particularly direct and pointed contribution to this counteroffensive.<sup>3</sup> Like all of Cassam's work, it contains much careful philosophical argumentation; but it aims to be a book, not just for philosophers, but for any thinking person interested in self-knowledge. The discussion is admirably engaging and down-to-earth throughout.

One major aim of the book is to challenge the idea that our mode of awareness of our own attitudes differs in principle from our mode of awareness of the attitudes of others. Cassam grants that we may in fact know more than other people do about our own present beliefs, desires, and intentions, at least when these attitudes concern matters that are not bound up with our self-conception; but he denies that any such *de facto* asymmetry in how much we know reflects a fundamentally different mode of awareness. I discern the attitudes of other people by a kind of interpretation: I draw inferences from their words, deeds, and other expressive behavior. When I ascribe attitudes to myself, my evidence typically includes more than this, since I have access not only to such outward signs but also to my own inner monologue and whatever images or impressions may accompany it. Nevertheless, Cassam argues, I pass from this evidence to knowledge of my own attitudes by taking a step of the same broadly *inferential* kind as the one I take in interpreting others. The relevant inferences may occur effortlessly and automatically, and this may give us the impression that we have immediate and noninferential access our own minds; but from an epistemological perspective, attitudinal self-knowledge is no less inferential than knowledge of the attitudes of others.

The larger aim of *Self-Knowledge for Humans*, however, is not merely to defend such inferentialism. As his title indicates, Cassam thinks mainstream work on self-knowledge fails to offer an account of self-knowledge ‘for humans’. The sort of humanity Cassam has in mind is not the sort that elevates us above the beasts, but the kind evoked by the phrase ‘only human’, as in the Human League’s chart-topping single of 1986:

I’m only human  
Of flesh and blood I’m made  
Human  
Born to make mistakes

Humans in this sense stand opposed to Vulcans, that remarkable race of alien beings whose rationality is never hampered by haste, emotion, fatigue, or self-deceit. That we *homo sapiens* are not such beings is widely recognized – the Human League were aware of it, and it is a crucial premise of *Star Trek* – but Cassam points out that our anecdotal evidence for this conclusion has been powerfully reinforced in the recent decades by rigorous studies of human reasoning and decision-making.<sup>4</sup> We are, it turns out, systematically liable to form beliefs and make choices, not on the basis of rationally defensible principles, but by employing simplifying ‘heuristics’ which get us to the optimal outcome in some cases, but lead us astray in others. Moreover, we typically do not recognize the role that such heuristics play in our thinking, and often are disposed to give demonstrably confabulated accounts of our reasons for believing or choosing as we do.

All this is certainly discouraging news, but who should be surprised by it? According to Cassam, it should come as a surprise to a group of philosophers he calls ‘Rationalists about self-knowledge’, whose core idea is that we can know our own minds by employing the ‘Transparency Method’, which tells us that we can determine our own attitude toward a given proposition *p* by considering some corresponding first-order question about *p* itself. Cassam argues that philosophers who think this strategy can provide a general account of how we know our own attitudes are committed to supposing that, in general, our attitudes are as they rationally ought to be. If this is right, then the evidence that humans are only imperfectly rational constitutes *prima facie* evidence that the Rationalist theory of self-knowledge does not apply to humans, but to some other superhuman sort of creature. Cassam calls this other sort of creature ‘homo philosophicus’, by analogy with the idealized ‘homo economicus’ of classical economic theory. I will simply call them ‘Vulcans’.

Cassam sees Rationalism as epitomizing the philosophical tendency to ignore our humanity in theorizing about self-knowledge, but he thinks this charge applies in a broader sense to much mainstream discussion of the topic. In general, he holds that philosophers have exaggerated both the extent to which we have privileged knowledge of our own minds and the immediacy with which we have this knowledge when we do. Indeed, he suggests that mainstream philosophical work on self-knowledge has lost touch with the human interest of the topic. When ordinary people hear the term 'self-knowledge', they think of knowledge whose acquisition would be a substantial achievement: knowledge of my true character, for instance, or of which of my professed values I genuinely hold. Philosophers, by contrast, tend to focus on varieties of self-knowledge whose triviality would astonish ordinary people: knowledge of whether I believe it is raining, for instance, and knowledge of whether I presently feel toothache. Why do philosophers engage in scholastic disputes about such matters when there are real and significant questions that beg to be addressed? This, in effect, is Cassam's challenge to mainstream philosophical work on self-knowledge.

I think this is a valuable challenge: it *is* striking that so much philosophical discussion is devoted to relatively trivial self-knowledge. What justifies this focus? I hope to say something about this, but in order to do so, it will be necessary first to address some misrepresentations in Cassam's presentation of the transparency approach and to raise some doubts about whether we can be satisfied with Cassam's inferentialism. These tasks will occupy me in the next two sections. In the final section, I will return to the question of the philosophical (and human) interest of 'trivial' self-knowledge.

#### 1. Transparency and Vulcanism

Cassam's book is cast as a rebuke to mainstream philosophical theorizing about self-knowledge as a whole, but it is particularly addressed to approaches that emphasize the 'transparency' of questions about our own attitudes to questions about the world toward which our attitudes are directed. These approaches, which have loomed large in recent discussion, take their departure from the observation that, in various cases, there is a correlation between how I *would* answer some world-directed question and how I *should* answer some question about my own attitude. Given this correlation, they argue, I can answer the question whether I hold a given attitude by treating it as 'transparent' to a corresponding world-directed question. The most widely discussed case of transparency concerns belief: Gareth Evans (1982), Richard Moran (2001), and

others have argued that I can determine whether I believe that  $p$  simply by treating this question as transparent to the question whether  $p$  itself is true. When consideration of the latter question leads me to affirm that  $p$ , I should affirm that I believe that  $p$ ; otherwise not.

Cassam contends that this observation could supply the model for a general account of attitudinal self-knowledge only if we were Vulcans. But why should this be? There is no need to mention rationality in stating the basic idea of the transparency approach, at least as I have characterized it. And what I have just said about the transparency of belief is perfectly standard, and fits the characterizations Cassam himself quotes (cf. 3, 102).<sup>5</sup> To generalize the approach, it would be necessary to find world-directed questions whose answers correlate with other attitudes. But what feature of this project requires us to presuppose that, in general, we hold only attitudes that we rationally ought to hold?

Cassam makes a three-step case for the connection between the transparency approach and the idea that we are Vulcans. First, he notes that, at least according to Richard Moran's influential account, doxastic transparency is intelligible only insofar as I am entitled to assume that what I believe regarding  $p$  is 'determined by my reflection on [the] reasons' that speak in favor of  $p$  (Moran 2003: 405; quoted at Cassam 2014: 3). Not all transparency theorists agree with Moran about this, but there is at any rate a significant body of work (including work by the present reviewer) that accepts Moran's claim about the link between transparency and reflection on reasons. Second, Cassam observes that it is not clear how we might generalize Moran's idea to attitudes other than belief: what sort of reflection on reasons would allow me to determine what I fear, or want, or hope? Moran does not discuss this problem in detail, but Cassam notes that David Finkelstein has proposed a 'helpful recasting' of Moran's idea that looks like it might admit of generalization. According to Finkelstein,

The question of whether I believe that  $P$  is, for me, transparent to the question of what I ought rationally to believe. (2012: 103; quoted at Cassam 2014: 4)

In general, Cassam suggests, we may take transparency theorists to hold that I can treat the question whether I hold some attitude  $A$  as transparent to the question whether I ought rationally to hold  $A$ . Finally, Cassam suggests that this 'Transparency Method' (TM) for determining my own attitudes will reliably lead me to knowledge only on two assumptions: first, that my attitudes are in general as they rationally ought to be; and second, that I am entitled to assume this about myself (4-5, 51-2). The first assumption is that we are Vulcans, or near enough; the second is that we are entitled to believe ourselves to be Vulcans.

Much of Cassam's book is then devoted to arguing that we are not Vulcans. He also objects that, in many cases, it may be perfectly clear to me what my attitude on a given topic is, although it is quite unclear what attitude, if any, I rationally ought to hold on that topic (6-7, 104-6). Finally, he suggests that, even where it is possible to know one's own attitudes by TM, this is itself a case of inference, mediated by the premise that one's attitude on a given topic is as it rationally ought to be (6, 111-12, 117-18).

All of these objections rest, however, on the cogency of Cassam's case for attributing to transparency theorists a commitment to supposing that our attitudes are as they rationally ought to be. I think the case is flawed at several points. The most fundamental issue concerns what Moran means when he claims that doxastic transparency is intelligible only if I can assume my belief regarding *p* to be determined by my reflection on reasons in favor of *p*, but let me set this aside for a moment and address the latter two steps.

First, a word about generalizing the transparency approach beyond the case of belief. I do not accept Finkelstein's proposal as a charitable interpretation of the transparency approach. Cassam's book might give one the impression that Finkelstein is an advocate of this approach, but in fact he is a critic aiming to make it out to be silly. It would indeed be silly to propose that, in general, I can know whether I hold some attitude *A* by considering whether I ought rationally to hold *A*, but this is simply not the view of Moran or other major advocates of this approach.<sup>6</sup> In the case of belief, as we have seen, the world-directed question to which the question about whether I believe that *p* is transparent can be stated without any mention of rationality: it is simply the question whether *p*. To generalize the approach, it would be necessary to identify world-directed questions corresponding to other attitudes, and there has been a good deal of work on this, which Cassam does not review. Evans suggests, for instance, that whether it visually appears to me that *p* is correlated with whether *p* would figure in my answer to the question how things are directly in front of me, when I constrain my answer in certain specifiable ways (Evans 1982: 227-8). Alex Byrne (2011) has argued that whether I intend to do *A* is correlated with my answer to the question whether I will do *A*, when I answer the latter question only on a certain sort of basis. Characterizing the relevant world-directed questions is not a simple task – not significantly simpler than, and closely related to, the task of giving a philosophical analysis of the relevant attitudes. But the idea that such characterizations can in principle be given is no less plausible than the idea that, to each type of attitude, there corresponds a characteristic stance on some question about the non-mental world. And this is

surely an interesting idea, and not obviously a false one. At any rate, no reason has been given why implementing this suggestion requires appealing to questions about what attitudes one *ought rationally* to have.

Let us, however, grant for the sake of argument that TM says I should treat the question whether I hold a given attitude A as transparent to the question whether I ought rationally to hold A. *Even if this is granted, it does not follow that the soundness of TM depends on my attitudes being in general as they rationally ought to be.* Let us focus for simplicity on the case of belief – the same point could be made for any other attitude. Suppose it is the case that my belief about some proposition *p* is not, as a matter of fact, as it rationally ought to be: no matter how carefully I consider the overwhelming evidence that not-*p*, still I find myself returning to the view that *p*. Following TM could *still* lead me to reliably correct answers about whether I believe that *p* provided that the same imperviousness to evidence affects my judgment about whether I rationally ought to believe that *p*. To raise a problem for Cassam's argument, I need not insist that this is in fact regularly the case (though I suspect it may be): it suffices that TM *could* be a reliable method for determining one's attitudes even if one's attitudes were not in general as they rationally ought to be. If this is right, then someone who claims we can know our own attitudes by TM is not thereby committed to assuming that our attitudes are generally as they rationally ought to be.<sup>7</sup> There is simply no direct connection between TM and the assumption that we are Vulcans, even on Cassam's contentious understanding of TM.

What about Moran's claim that doxastic transparency is intelligible only if I can assume my belief concerning *p* to be determined by my reflection on reasons in favor of *p*? Doesn't this posit some sort of connection between transparency and rationality? Yes, but a quite different sort from the one Cassam supposes. Transparency theorists begin by noting a connection between the self-directed question whether I believe that *p* and the world-directed question whether *p*. The latter question invites, not an investigation of my own psychological state, but an assessment of whether the relevant worldly proposition is true: it calls on me to exercise my power to *judge* whether *p*.<sup>8</sup> On some occasions, I do not need to deliberate about this question: my view is already settled. On others, I may consider evidence but assess it in a way that deviates from norms of rationality: perhaps my judgment results from the operation of one of the natural but unsound heuristics identified by psychologists. Regardless, the power I exercise in answering the question whether *p* is my power to judge whether *p*. But no one could have this power without having the capacity to consider evidence for and against the truth of a

proposition, and to affirm or reject propositions on the basis of such consideration. The question I confront in exercising this power is, as Moran puts it, a 'deliberative question': a question about whether *to* accept the relevant proposition. The infinitival form of the verb here indicates the characteristic perspective of a rational capacity: it regards my stance on whether *p* not as a given fact, but as mine to settle. This holds true even if I answer by reaffirming an already settled view, or arrive at my judgment in a way that is normatively defective.<sup>9</sup>

When Moran speaks of assuming that my belief concerning *p* is 'determined by my reflection on reasons' in favor of *p*, what he has in mind, I think, is just this: that in order for it to make sense for me to treat the question whether I believe that *p* as transparent to the question whether *p*, I must presume that my belief as to whether *p* just is the answer I am disposed to reach by exercising my rational capacity to judge whether *p*.

It is, of course, possible for a person's reflective judgment on *p* to part company with her underlying belief on the matter. When she explicitly considers whether *p*, she may reach one conclusion, but another view may be reflected in her persisting tendencies to react and respond. In this case, TM will not offer her a sound route by which to assess whether she believes that *p*. But this is just Moran's point: a person whose reflection on whether *p* parts ways with her underlying belief cannot know her belief simply by looking out toward the world; she must observe herself and draw inferences. Thus, where transparency does not obtain, we can say that a person is *alienated* from her own belief: she is not capable of knowing it from a participant's standpoint, so to speak. Where transparency does obtain, by contrast, a person's belief is precisely what is expressed in her rational reflection. But when I speak of 'her rational reflection' here, I am just referring to her exercise of her capacity explicitly to consider a question and make a judgment. She may well exercise this capacity in a way that is not 'rational' in the normative sense.

Indeed, the connection between self-knowledge and reflection on reasons can hold even in the case of 'recalcitrant attitudes' that are impervious to our all-things-considered judgments. If I have an unconquerable fear of flying that persists even as I judge there to be no real basis for fear, still, insofar as I am not alienated from my fear, I will be able to see the world from its perspective, so to speak. It is not just that, when I fly, I will find my brow sweating and my hands gripping the armrests; I will experience flight as perilous, and have my reasons for concern: this moment of turbulence; that unexplained knocking noise. My fear has its own rational perspective on the world, and I can give myself over to this perspective, and know that I



am afraid, and what I am afraid of, in this way. So, contrary to what Cassam supposes (107ff.), a version of Moran's transparency point applies even to recalcitrant attitudes. The point concerns the connection between attitudinal self-knowledge and our rational capacity for judgment. It simply has no direct connection with the assumption that we are Vulcans.

Transparency theorists are thus not committed to supposing that we determine what our attitudes are by asking what attitudes we rationally ought to hold; and even if they did maintain this, they would not be committed to assuming that we are Vulcans. I conclude that 'Rationalism about self-knowledge' is a chimera constructed by fusing the transparency approach to a commitment that transparency theorists have no reason to accept. Like classical chimeras, this is a mythical creature.

## 2. Inference and Alienation

Cassam's account of how we know our own attitudes is that, quite generally, we know them by inference. To illustrate what he has in mind, he invokes a case described by Krista Lawlor (2009), in which a woman considers whether she wants another child and comes to recognize that she does by paying attention to her own thoughts and experiences: a dreamy state of recollection she falls into while boxing up her first child's outgrown clothes, a feeling of envy that comes over her at the news of a friend's pregnancy, and so on. Following Lawlor, Cassam describes this woman as making an 'inference from internal promptings' (143): she infers her standing attitude of wanting another child as the best explanation of certain conscious episodes, against the background of some tacitly-held theory of how wants express themselves.

Moreover, Cassam suggests that this sort of inferential account applies not only to cases like Lawlor's, but even to the kinds of cases favored by transparency theorists, such as knowing my belief concerning  $p$  by considering whether  $p$  and making a judgment. For, Cassam argues, judging that  $p$  is just one more conscious episode, which may be good evidence that I believe that  $p$ , but only on the assumption that, at least in this case, my conscious judgment reflects my standing belief (111, 117).<sup>10</sup> So this too is an inference, even if it occurs effortlessly and automatically. Indeed, Cassam argues, since it is not plausible to hold that we know our own attitudes by inner perception, and not acceptable to claim that we know them 'from nothing', inferentialism is really 'the only game in town' (158).

In a fuller discussion I would want to contest several points in this argument, but in this review I will focus on just one worry about Cassam's inferentialism: that it would imply that, in

general, we are alienated from our own attitudes, whereas it is clear that normally we are not alienated from them. This is Richard Moran's objection<sup>11</sup> to inferentialism, and Cassam is aware of it, but he does not think it succeeds. His reply is that, if to be alienated from an attitude is to find it persisting even though I reject it, or do not identify with it as expressing my view of the world, then the fact that I have inferred my attitude from inner promptings need not imply alienation. I *may* be alienated from some of the attitudes I know by inference, but there is no reason why I *must* be alienated from all of them (157).

I think this misses the point. Even if I endorse the attitudes I ascribe to myself, there is a kind of alienation implied in the idea that I should need to discover them by inference. Consider a case of knowing that I believe that *p* by judging that *p*, and suppose for the sake of argument that Cassam's analysis of this case is sound. I consider whether *p* and reach the judgment that *p*. So some part of me has taken an affirmative stance toward *p* – call him 'MB *qua* judger'. But by hypothesis, there is another part of me, 'MB *qua* believer', whose stance on *p* might yet differ. We seem to have two points of view here: one which accepts that *p*, another which may or may not accept it. Now, such a division can arise within a person: what I judge when I consciously consider whether *p* may come apart from the attitude toward *p* reflected more broadly in my behavior. But Cassam's inferentialism builds such a division into its characterization of even the most unproblematic case of attitudinal self-knowledge. It does not matter that the stance on *p* expressed in judgment may for the most part *coincide* with the stance embodied in belief. The point is that, even when the judging subject has settled on an answer to the question whether *p*, there is still (in principle) another question for her to settle, namely whether she actually believes that *p*.

Many contemporary philosophers are committed to this sort of picture of the epistemic relation between judgment and belief, but I think on reflection it should strike us as strange. Focusing on a case in which the subject judges out loud may help to bring out its strangeness.<sup>12</sup> Let the subject consider whether *p* and express her conclusion aloud:

(1) Yes, *p*.

Now let her infer her own belief from this 'external prompting', and again express her conclusion aloud:

(2) I believe that *p*.

Assuming the subject is not alienated from her own belief, we would ordinarily take (2), like (1), to express *belief that p*. On the inferentialist analysis, however, this cannot be right. Whereas

(1) may express belief that  $p$ , (2) relates to this belief only indirectly: it merely expresses the subject's belief *that she believes that  $p$* , which is based on evidence that she believes that  $p$ , but is not itself an unmediated expression of her first-order belief. So her assertion of (2) expresses, not an endorsement of  $p$  as true, but a (no doubt very well founded) hypothesis about herself. If she were fully clear about her own epistemic situation, she really ought to say to herself: 'Yes,  $p$ , and so it is extremely likely that I believe that  $p$ , since my beliefs coincide with my judgment in most cases, and I've no reason to think this is an exception.' If I received this sort of report on someone's beliefs, I should demand to speak to the believer herself, not just to her biographer, however well-informed she might be.

The uncanniness of the inferentialist analysis is even more palpable in the case of other attitudes. Consider hope. Suppose the subject notices inner promptings (whatever these might be) that she takes to reflect a hope that  $X$  will come with her on a picnic. Let her infer that she hopes  $X$  will come and express her conclusion by saying to  $X$

(3) I hope you'll come.

If (3) is offered as a report on an inference, I think  $X$  should find it rather disappointing. 'I hope you'll come' is normally heard as welcoming because it is taken to express hope itself, not just belief, however certain, in the existence of hope.<sup>13</sup> Even if the subject adds: 'And I (*qua* judger) wholly endorse and identify with this attitude', her remark is still missing its marrow: the hope itself is elsewhere, and we are merely looking on. Parallel points could be made about other attitudes. To treat all attitudinal self-knowledge as inferential is to build alienation into the very structure of such knowledge.

Indeed, even in the case described by Lawlor, I think inferentialism gives an unnatural account of the situation. A person can certainly realize that she wants another child by paying attention to her own thoughts and feelings in the way Lawlor describes, but is it really plausible to represent this as a matter of detecting some standing fact of the matter? Her feelings when she boxes up outgrown clothes and receives news of her friend's pregnancy are certainly indications of an incipient desire, but 'incipient' is important here. It is natural to imagine her also thinking of ways in which having another child would make it difficult to pursue other things she cares about. What she wants to know, presumably, is whether the decision to have another child is one she can genuinely embrace, and though 'inner promptings' may serve as indications of such a readiness, this is not simply a question of discovering what is already so but of

reaching a settled attitude on the matter. To investigate this as if it were a matter for discovery on the basis of evidence sounds, even here, like alienation, or indeed like bad faith.

But what is the alternative? Consider once again the simple case in which I reflect on whether *p* and judge that *p*. Mustn't I still make an inference to conclude that I believe *p*? And doesn't this inference rest on the premise that my judgment reflects my standing belief, either in general or at least in this case? This is not the place for a full treatment of this topic, but let me make two points. First, it is undeniable that, in unfavorable cases, I can *seem to myself* to judge that *p* although I don't believe that *p*, but it is contentious to describe these as cases in which I *judge* that *p* without believing that *p*. To describe them in this way is to define judgment so as to ensure there is only a contingent connection between judging and believing, and then it looks inevitable that, even in favorable cases, my judging that *p* can't justify me in ascribing a corresponding belief to myself without some further premise. But it should not be granted without argument that, even in favorable cases, my judging that *p* leaves open the possibility that I may not have the relevant belief. The inference from my epistemic situation in unfavorable cases to my epistemic situation in favorable cases is no more straightforward here than in the philosophy of perception, and in the latter domain a significant number of philosophers reject the inference from the fact that, in unfavorable cases, it can perceptually seem to me as if *p* when it is not the case that *p*, to the conclusion that, even in favorable cases, my perceptual basis for believing that *p* leaves open the possibility that not-*p*. Likewise, it might be argued that, in favorable cases, judging that *p* gives me a kind of justification for self-ascribing a belief that does not leave open the possibility that I do not believe that *p*, hence a justification that needs no supplementation by a further premise.

Secondly, the idea that I must draw an *inference* from my judgment that *p* to the existence of a corresponding belief looks most compelling if we assume that what warrants me in ascribing a belief to myself is my awareness *that I have made a certain judgment*. But again, this way of representing the matter is contentious. According to transparency theorists like Evans and Moran, I can move directly from answering the world-directed question whether *p* to answering the self-directed question whether I believe that *p*. Their idea, I take it, is not that an awareness of *my own judgment* must mediate this step; it can suffice that I am (at least seemingly) aware *that p*, without any need for a distinct step of recognizing *that I have so judged*. I have tried to defend this idea elsewhere.<sup>14</sup> Here I will simply remark that Cassam's interpretation – that what justifies me in ascribing a belief to myself is an awareness of my own

judging – seems to me already to introduce a dissociation between the perspective from which I judge about the world and the perspective from which I self-ascribe beliefs. On the resulting view, the subject is not in a position simply to look outward and affirm: ‘I believe that  $p$ ’; her basis for this self-ascription is rather an inner event whose occurrence she monitors. I believe this picture is not compulsory, and not attractive once we appreciate its significance.

### 3. Self-Knowledge and the First Person Perspective

Let me conclude by returning to Cassam’s challenge to mainstream theorists of self-knowledge: that they have focused on trivial self-knowledge and thereby lost touch with the human interest of their topic. Can something be said to justify all the ink philosophers have spilled discussing how one knows that one believes it is raining?

Cassam’s challenge rests in part on his claim that philosophers have exaggerated the epistemological distinctiveness of attitudinal self-knowledge, and his case for this depends on his inferentialism, which I have already disputed. There is, however, a way of pressing Cassam’s question that does not require the truth of inferentialism. Much of the force of his challenge comes simply from the observation that attitudinal self-knowledge is most readily available where the attitudes in question have the least personal significance (cf. 30, 110). Where my attitude on a given topic matters to my self-image, it is common enough for my own assessment of it to be unreliable. In the face of this, how can we justify focusing exclusively on cases of easy attitudinal self-knowledge?

I think our observations about transparency and alienation provide the basis for an answer this challenge, but let me emphasize that I will not try to justify inattention to the kinds of self-knowledge that are hard to achieve. I agree that it is worth understanding what difficulties stand in the way of substantial self-knowledge, and how such knowledge matters to a good life. It is, I think, not easy to frame questions about these topics that philosophers are well-equipped to address, and perhaps this accounts for the relative paucity of attention they have received. But I certainly do not want to defend this neglect, just to defend the interest of the topic that philosophers have more commonly discussed under the heading of ‘self-knowledge’.

Cassam anticipates that, if philosophers are asked to justify their focus on easy self-knowledge, they will respond that this knowledge is epistemologically distinctive in a way that matters to philosophy, even if it matters little to humanity at large (43ff.). This may well be the

most common response to such a challenge, but if so, I think the majority opinion does not reflect the best reasons for interest in this topic. A more compelling answer is implicit in Sydney Shoemaker's remark that

it is essential to a philosophical understanding of the mental that we appreciate that there is a first person perspective on it, a distinctive way mental states present themselves to the subjects whose states they are, and that an essential part of the philosophical task is to give an account of mind which makes intelligible the perspective mental subjects have on their own mental lives.  
(Shoemaker 1996: 157)

In recent decades, philosophers interested in the 'distinctive way mental states present themselves to the subjects whose states they are' have mainly focused on the 'phenomenal' aspect of conscious mentality, but I take Shoemaker to hold that the special way in which we are aware of our own attitudes is another crucial aspect what he calls 'the first person perspective' on mind.<sup>15</sup> His claim is that an understanding of this perspective is essential to an adequate understanding of the mind itself.

I cannot defend this claim here; there is only time to express sympathy for it, and to suggest that it would provide a powerful rationale for interest in the awareness we have of (certain of) our mental states in virtue of being the subject of those states. Note that this rationale addresses us, not primarily as epistemologists, but as philosophers of mind: the idea is that, to understand the mind, we must understand subjectivity, and subjectivity is expressed primarily in a special mode of awareness of certain states: awareness of them from a special standpoint one has precisely in virtue of being in those states.<sup>16</sup>

Our observations in the foregoing sections should begin to give us some grip on how this idea might apply to attitudinal self-knowledge. We have seen that the capacity of a subject to answer a question about her own attitude on some topic by treating it as 'transparent' to a question about the world reflects a special kind of participant's standpoint on that attitude, a standpoint that contrasts with an alienated standpoint in which the subject merely knows *of* the attitude as a fact about herself. I believe understanding knowledge of an attitude *from a participant's standpoint*, and how it differs from mere cognizance (as we can call it), is crucial to understanding what it is for reflective creatures like us to have a standpoint on the world; and this is surely a core element of the concept of mind, whatever else may be contained in it.

To assert this sort of essential connection between mentality and subjectivity is not to imply that a subject will always know her own mind, any more than the essential connection between, say, the heart and the circulation of the blood ensures that the heart will always

perform its function: in each case, there is much that can interfere. But if there is such a connection, there will be a point in studying cases where things do *not* interfere: viz., cases of easy self-knowledge. To my mind, at least, this is the best rationale for studying the kind of self-knowledge that Cassam calls 'trivial'.

## References

- Boyle, M. (2011), 'Transparent Self-Knowledge.' *Proceedings of the Aristotelian Society*, Supp. Vol. 85: 223-41.
- Burge, T. (1996), 'Our Entitlement to Self-Knowledge'. *Proceedings of the Aristotelian Society* 96: 91-116.
- Byrne, A. (2011), 'Transparency, Belief, Intention'. *Proceedings of the Aristotelian Society*, Supp. Vol. 85: 201-221.
- Carruthers, P. (2011), *The Opacity of Mind*. Oxford: Oxford University Press.
- Cassam, Q., ed. (1996), *Self-Knowledge*. Oxford: Oxford University Press.
- Cassam, Q. (2014), *Self-Knowledge for Humans*. Oxford: Oxford University Press.
- Evans, G. (1982), *The Varieties of Reference*. Oxford: Oxford University Press.
- Finkelstein, D. (2012), 'From Transparency to Expressivism'. In J. Conant and G. Abel, eds., *Rethinking Epistemology*, vol. 2. Berlin: De Gruyter.
- Kahneman, D. (2011), *Thinking, Fast and Slow*. London: Allen Lane.
- Lawlor, K. (2009), 'Knowing What One Wants'. *Philosophy and Phenomenological Research* 79: 47-75.
- Moran, R. (2001), *Authority and Estrangement*. Princeton: Princeton University Press.
- \_\_\_\_\_. (2003), 'Responses to O'Brien and Shoemaker'. *European Journal of Philosophy* 11: 402-19.
- Ross, L. and Nisbett, R. (2011), *The Person and the Situation: Perspectives of Social Psychology*. London: Pinter & Martin.
- Schwitzgebel, E. (2011), *Perplexities of Consciousness*. Cambridge, MA: MIT Press.
- Shoemaker, S. (1968), 'Self-Reference and Self-Awareness'. *Journal of Philosophy* 65 (19): 555-67.
- \_\_\_\_\_. (1996), *The First-Person Perspective and Other Essays*. Cambridge: Cambridge University Press.
- Wilson, T. D. (2002), *Strangers to Ourselves: Discovering the Adaptive Unconscious*. Cambridge, MA: Harvard University Press.

---

<sup>1</sup> An earlier version of this essay was presented at an author-meets-critics session at the 2015 Pacific Division Meeting of the American Philosophical Association in Vancouver, Canada. I am indebted to



---

Richard Moran and Ram Neta for comments on a draft, to Quassim Cassam for responses when the paper was presented, and to several helpful comments from the audience.

<sup>2</sup> A representative sample of work in this tradition can be found in Cassam 1996.

<sup>3</sup> Other notable contributions include Carruthers 2011 and Schwitzgebel 2011.

<sup>4</sup> Cassam relies principally on the well-known work of Kahneman (2011), Ross and Nisbett (2011), and Wilson (2002).

<sup>5</sup> Except where otherwise indicated, all page references are to Cassam 2014.

<sup>6</sup> Moran does speak of the question whether I believe that *p* being transparent to the question whether *to believe* that *p*, but he distinguishes this ‘deliberative question’ from the ‘normative question’ whether I *ought to believe* that *p* (Moran 2001: 59, 145-6). I say more about how I understand Moran’s ‘whether to believe’ formulation below.

<sup>7</sup> Cassam in fact acknowledges that, in some cases, a subject whose attitudes are normatively irrational may nevertheless be able to determine what attitudes she holds by using TM (cf. 107), but he does not seem to recognize the problem this poses for his claim about the connection between TM and the assumption that we are Vulcans.

<sup>8</sup> N.B.: Not to judge whether I ought rationally to believe that *p*. I do not ask *whether it would be rational for me to believe that p*, but *bring to bear my capacities for rationality* on the question whether *p*.

<sup>9</sup> Cf. Moran 2001: 148.

<sup>10</sup> Much of Cassam’s discussion is conducted on the assumption that the relevant judgment is: *I ought rationally to believe that p*. I have already contested this way of representing the transparency proposal. Here I set aside this feature of Cassam’s view. A version of Cassam’s case for inferentialism can be made without this assumption (as Cassam recognizes: cf. 117).

<sup>11</sup> See Moran 2001, esp. Chapter 4. For related points, see Burge 1996.

<sup>12</sup> I consider verbal expressions of thought for vividness, but I do not think anything turns on this. The subject’s private thoughts about her own attitudes would stand in an equally uncanny relation to her attitudes; imagining these thoughts uttered in a conversational situation just focuses our attention on what is missing.

<sup>13</sup> Cf. Moran 2001: 98ff. Note that my claim that the assertion of (3) normally expresses hope does not rule out that it is a *comprehending* expression of hope, so that the expression includes, inextricably, an expression of the subject’s awareness of hoping. It is only given the inferentialist analysis that (3) must express belief that one hopes *rather than* hope itself.

<sup>14</sup> See Boyle 2011.

<sup>15</sup> An indication of this is that Shoemaker cites the remark just quoted as a way of explaining what ties together the essays collected in Shoemaker 1996, a volume that centrally includes his well-known papers on attitudinal self-knowledge.

<sup>16</sup> I believe there is also an important connection, not considered by Cassam, between the specially immediate awareness we can have of certain of our own mental states and the question what it is to think of oneself ‘first personally’. Oedipus knew plenty of facts about the person who was in fact himself, but only at the denouement of *Oedipus Rex* did he understand that these were facts about *himself*. What is involved in understanding this (i.e., in grasping propositions of the form *I am F*)? There are powerful reasons for thinking that an answer to this question must connect first person thought with immediate self-knowledge. For classic discussions, see Shoemaker 1968 and Evans 1982, Chapter 7.